PedScope
User Guide

Version 2.3.01

March 2014

Table of Contents

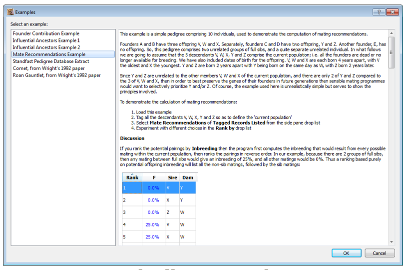
Getting Started

Begin by selecting **PedScope** from the Windows **Start** menu, or from an icon on your desktop.

To load data, choose **File|Open** (i.e., select the **Open** option on the **File** menu). This accepts text/CSV files in a variety of formats. For further information see Importing Data.

You might like to know that PedScope has various examples built in for demonstration purposes. To choose one of these select **File|Examples**. You can also create a simulated pedigree, useful if you want to create a large data set just to explore the program's handling of large pedigrees - choose **File|Simulate Pedigree**.

Once the file is loaded the Main Window is displayed.



Inbuilt Examples

Importing Data

To load data, choose **File|Open**, then select your import file. The file needs to be a text/CSV file - the kind of file that you might export from a spreadsheet - but which can also be produced by many database and pedigree programs.

You are then taken through a series of windows that let you set options to control the way PedScope processes the import file. ***Note that in many cases there will be few if any of these settings that you need to change***, because PedScope inspects the data file and chooses sensible defaults for the file format and its contents.

NOTE: To see some examples of acceptable import data see: [Import Examples].

Character Encoding

The **Character Encoding** window selects the character set encoding and line ending format. If the options presented here don't mean anything to you, ***don't worry because in practice you are very unlikely to need to change these settings***. If in doubt simply click **Next** to proceed.

The **character set** defines how the bytes in the file are encoded, such as UTF or ISO-8859-1. The **line ending format** says whether input lines are terminated by carriage return, line feed, or both.

Column Formatting

The **Column Formatting** window controls the way that the columns are formatted in the file. This is where you choose whether the file is comma or TAB separated, whether there is a header line, and whether column values are quoted. Again, it is unlikely you will need to change anything here. Typically you just need to click **Next** to proceed.

Columns

The **Columns** window is where you define the meaning of each import column - i.e. where the name, sex, sire etc. are to be found in the input data. It is more likely that you will need to alter the settings here than in the previous windows.

However, in many cases again PedScope is still able to deduce the correct meaning of columns in your data file. Before it displays the **Columns** window, it will inspect the first few lines of the input file and attempt to make a reasonable guess as to which columns contain the main pedigree data - name/id, sire, dam, sex and date of birth.

To make changes, click the **Customize** checkbox. It can help at this stage to make the window wider and deeper. The rows shown are as follows:

Row	Purpose
Field	The meaning of the data in this column.
Type	The type of the data, such as Text, Integer, Date, etc. For some fields, e.g. date of birth, the type is implicit and you cannot change it.
Label	The label that is to be used for the field once the data has been imported into PedScope. This is only meaningful for fields that will become user fields in the main table.
Input Header	This will be shown in blue and shows the text from any header line that was found in the input data file. This is displayed purely for your convenience when checking that the Field and Type have been set correctly.
Input Data	Up to 4 lines of data, again shown in blue, taken from the input data file, for reference.

You should review the **Field**, **Type** and **Label** settings for each input column to be sure they are correct for the particular input data. Ensure that the date format and male/female sex values are consistent for the input data file, then click **Finish** to proceed with the import.

Fields

The full list of fields is given in the table below.

NOTE: Several of these fields are related to each other and need to be consistent within the input file in order to successfully import data. These are the **Name** and **ID** fields.

Each input record (i.e., each animal) must have an ID field, a Sire ID field and a Dam ID field. These are used by PedScope to construct the ancestral linkages - i.e. they are how it works out which parents any given animal is descended from. The ID fields can be numeric or text. Numeric IDs are generally record numbers/keys as generated by whatever software created the import file.

Optionally, the import data may also contain a **Name** field.

If you are importing data that contains just the name of each animal, its sire and dam, without any kind of record number/key, you should import these columns as the Animal ID, Sire ID and Dam ID (and not as 'Name' fields).

Obviously the Sire/Dam ID/Name fields need to contain consistent data so that PedScope can work out the ancestral links correctly.

Field	Description
Animal ID	A unique number or key for this animal. The Animal ID is mandatory if the Name is not given, otherwise optional.
Sire ID	Sire key/number. If present, the Animal ID field must also be included.
Dam ID	Dam key/number. If present, the Animal ID field must also be included.
Name	Animal name.
Sex	Gender. The allowed values for male/female are given in boxes beneath the import definition table.
Date of Birth	The date format is set using a dropdown menu beneath the import definition table.
Date of Death	The date of death uses the same date format as date of birth.
Sire Name	Name of the sire. If present, the Animal Name field must also be included.

Dam Name	Name of the dam. If present, the Animal Name field must also be included.
User Field	A user defined column. Such columns will be loaded and tabulated within the main record display. If you also specify a numeric type for such columns then PedScope will sort them appropriately, and can report statistics for such columns. E.g. if your data has a column that is some kind of sample measurement for each animal (say, weight), then if you set the column to be a user field with a numeric data type (see below) then PedScope will compute statistics such as the average weight (both for any tagged set of animals, but also for date period stats). You can also have it plotted in graphs, e.g. to correlate it against other input columns or computed values (such as inbreeding), etc.
Alive?	Living status - whether or not the animal is known to be living. If used, column values "TRUE", "T", "YES", "Y" or "1" are interpreted as meaning 'this animal is alive'.
Dead?	Dead status - whether or not the animal is known to be not living. If used, column values "TRUE", "T", "YES", "Y" or "1" are interpreted as meaning 'this animal is not alive'.
Ignore	A column to be completely ignored in the input file.

Data Types

The full list of data types is as follows:

Type	Description
Text	Simple text column.
Date	Date value. The date format is set using a dropdown menu beneath the import definition table.
Integer	An integer value (i.e. no decimal point).
Number	A number that may include a decimal point.

Examples

See Import Examples.

Import Examples

PedScope accepts input in various formats. For details on how to import data, and how to specify the input format, see Importing Data.

Shown below are various examples of different input data formats, all of which are acceptable. This is by no means exhaustive - the import window can cope with many different formats. For ease of display the examples shown here use comma separated text (CSV) but you could just as easily use TAB separated or other separator character.

Example 1: Very simple structure with a column for the gender and columns for the name, sire name and dam name.

```
SEX,NAME,SIRE,DAM
M,Comet,Favorite,Young Phoenix
M,Favorite,Bolingbroke,Phoenix
M,Bolingbroke,Foljambe,Young Strawberry
M,Foljambe,R. Barkers Bull,Haughton
F,Young Strawberry,Dalton Duke,Favorite cow
F,Phoenix,Foljambe,Favorite cow
F,Favorite cow,Alcocks Bull,
F,Young Phoenix,Favorite,Phoenix
```

Example 2: In this case the animal is given by both a name and an id number, and the sire/dam linkage is through the id number.

```
name,sire_id,dam_id,id
Arripay Blue Surprise,2,29,1,
Arripay Breaks Everything,26,22,2,
Arripay Brown Beauty,19,5,3,
Arripay Brown Boy,,,4,
Arripay Candy Girl,,,5,
Arripay Champagne Charlie,16,22,6,
Arripay Chocaholic,6,29,7,
Arripay Choccy Chips,,,8,
Arripay Dark Crystal,12,17,9,
Arripay Dark Magician,4,25,10,
...
```

Example 3: Similar to the above but this time with a column for date of birth. To control the format of dates (e.g. DD/MM/YYYY or MM/DD/YYYY etc) you can select from many supported date formats after the import window has been displayed.

```
ID,Sex,DOB,Sire,Dam
Abbots Wisdom,M,30/03/1930,Diver Of Woolley,Sewardstone Tess
Abbotsford Rosa of Tasvane,F,30/04/1986,Tasvane Taraius,Rachenco Sunshine
Alexander of Elsiville,M,15/05/1944,Torrdale Tinker,Zena Of Elsiville
Alexia of Tillwood Rachenco,F,30/04/1973,Spartan of Rachenco,Wayfarer of Tillwood
Alresford Advertiser,M,15/04/1951,Alexander of Elsiville,Alresford Mall
...
```

Main Window

The main window is divided into 1 or more sections. When you first open the a data file you will have, from left to right:

- The ***Main Record Table***. This lists the input data, with one animal per row. You should see columns for name/ID, sex, sire, dam and so on. The exact columns present depend not only on what is present in the input data but also the current document settings. It is possible to customize PedScope so that specific input columns are hidden, and to control which calculations are performed. For details of all the possible columns, see Main Table Columns.
- The ***Report Pane***. This displays one of several different reports/tables. Above the report pane is a pulldown menu which is used to change the displayed report type. We call this the ***report selector***. Initially the report selector is set to display the Relations report.
- The ***Property Pane***. This is used to show 'drill down' detail of whatever is currently selected in the report pane. The property pane is not always visible. Typically its used to show the record details for a highlighted result. E.g. with the Relations report, once you have relations listed - such as ancestors - you can select an ancestor (click it) and have all details for that ancestor shown in the properties pane.

You can resize the main window panes by dragging the divider between the panes.

Main Record Table

You can ***sort*** the main table by clicking any of the column headings. Clicking a heading again will reverse the sort order.

You will see some columns in blue. These represent values that PedScope has computed (you can change this color, and various other highlighting effects, by selecting **Preferences** from the **File** menu and clicking on the **Highlighting** tab).

Above the main record table are some pulldown menus to ***filter*** the displayed records. E.g. you can choose to only display animals that have offspring, or animals that are founders, or females only etc.

Below the main record table is a box labelled **Find**. To locate a record quickly e.g. find an animal given its name, first ensure the animals are sorted by name, then click in the **Find** box, enter the name, then click **Next**.

Near the "Find" controls are buttons labelled **Export** and **Print**. These are used to export the displayed table (e.g. to a spreadsheet or PDF file), and to print it.

Report Pane

The report pane - to the right of the main animal table - displays reports that are generated from one or more of the animals from the main table.

At the top of the report pane is a pulldown selector menu that you use to choose the type of report you want.

The way to choose the animals to which the report pane applies to is by using the pulldown menu that appears to the right of the report selector. This can be set to 1 of 3 options:

- **Selected Record** - the report is specific to a single animal, the one you most recently highlighted in the main record table e.g. by clicking a row.
- **Tagged Records Listed** - the report is specific to a group of one or more animals, the currently 'tagged' animals (see below).
- **All Records Listed** - the report is not specific to any of the animals listed, it will be 'run' for all animals listed.

Some reports are implicitly specific to a single record. E.g. the pedigree chart is specific to showing the pedigree for the selected animal, only. Other reports are only meaningful to multiple records, and for these you typically use tagging (see below) to select those records.

For further information on the various reports available please see:

- Relations

Tagging Records

Record tagging is the way to select multiple records for a report.

To tag a single row in the main record table: select it (click it or use 'Find' to locate it), then press CTRL+T (or choose **Record|Tag Selected Record**). When tagged the record is highlighted in the main table.

To tag a range of rows: select the starting row to be tagged and press CTRL+M (or choose **Record|Set Marker**). Then, move to the end of the range to be tagged, select that record by clicking it, and press CTRL+SHIFT+T (or choose **Record|Tag From Selection To Marker**).

When you use tagging to select the records for a report, make sure that you've set the report to be run for the **Tagged Records Listed** (you will find this option to the right of the report selector at the top of the report pane).

There are other options to control tagged records on the **Record** menu, e.g. for untagging.

There is also a **Tag All Matches** button below the main record table; this can be used to tag all records matching the 'Find' pattern.

There is a **Tagged Only** option above the main record table to limit it so it only shows tagged records (and possibly limited also to other record selection conditions you may have chosen, such as 'Males Only').

Tagged records are listed in the main table using a highlighted color. Different highlight colors can be set for males and females; by default these are a light blue and light pink. You can change these color, and various other highlighting effects, by selecting **Preferences** from the **File** menu and clicking on the **Highlighting** tab.

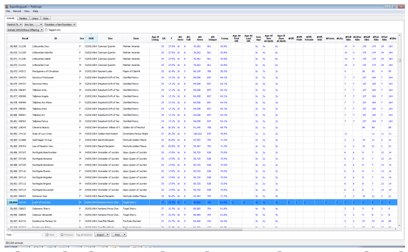
Other Record Types

Along the top of the main window there is a series of tabs normally labelled ***Animals***, ***Families***, ***Litters*** and ***Stats***.

You can use these to switch to display record types other than the animal records. These other record types are automatically generated from the input animal data by PedScope (you cannot import these directly):

- ***Families*** - PedScope groups animals together into what it calls 'families', where a family is defined as all animals that share the same parents. You can browse the families to see values and statistics computed at the family level, and easily browse all offspring of the family.
- ***Litters*** - PedScope further groups animals within individual families into 'litters', where a litter is defined as all animals that share the same parents and date of birth. The Litters tab is disabled if your input data does not contain data for the date/year of birth. Otherwise it operates similarly to the family records tab.
- ***Stats*** - last but by no means least we have the statistics tab. This is one of the most powerful features of PedScope. If your data includes date/year of birth, then PedScope will compute statistics for all animals falling in each period. You can control whether the periods are year, decade, month etc., and you can limit the stats to just the tagged records only. The program will then report summary stats on those animals born within the given date periods. The report pane will then let you drill down to give details on stats for individual periods, and plot graphs e.g. to plot average inbreeding over time. For further information see Statistics.

Main Table Columns



Main Window - Showing main record table with report pane hidden

This section describes the columns that can be displayed in the main record table.

Not all these columns are necessarily displayed; you can configure what is actually shown - see Customization.

For details on how to use the main table, including sorting, filtering and tagging, see Main Window.

The main table columns can be classified into 2 groups: those that are computed by analyzing the pedigree of each animal, and those that are not.

Columns That Do Not Require Pedigree Analysis

This includes all the main columns containing basic facts about each animal, such as name/ID, sex, parents, and so on.

Column	Symbol (if any)	Description
Record Name		Name of this animal, if it has one distinct from its ID.
ID		The input ID of this animal, if it has one that is distinct from its name.
Seq#		The input sequence number. This is always equal to the order in which the records were read in from the input file, with the first record given sequence number 1. This may or may not be the same as the ID or Rec#, depending on the order in which records are presented in the input data, and how your input IDs are assigned.
Rec#		The internal PedScope-assigned record number of this animal. This is not the same thing as the record name or ID as given in the input data, nor is it the same as the input sequence number. During the importing of data, PedScope has to renumber (reorder) the input records to ensure that they are in an order that guarantees that parents always precede their offspring. This ordering is required to implement certain algorithms quickly, and also enables PedScope to discover any self-parenting errors during input (a self-parenting error being when an animal appears to be descended from itself, which can happen when there are errors in the input data). See also the 'Age If Living' column (below).
Sire		Name or ID of the sire of this animal.
Dam		Name or ID of the dam of this animal.
Sex		The gender of this animal. Normally displayed as **M** or **F** (though you can change this: see Customization). It will be shown in a faded color if the sex has been deduced on input; this happens when the sex is not stated in the input data but has been deduced from the animal's appearance as the sire/dam of other animals.
Living*		Whether or not this animal is considered to be alive. This is determined from a combination of other information. Firstly, if the imported data has a 'Dead' or 'Alive' column value (see Importing Data), then this defines whether or not the 'Living' column is true. Failing that, if the date of death is known the 'Living' column is set to false. Otherwise, if the date of birth is known and the current age of the animal exceeds a user-chosen threshold, then again the 'Living' column is set to false. Finally, in the absence of any other way to determine the living status, it is set to a user-chosen default. See Customization.
Date of Birth	**DOB**	Date of birth of this animal or litter.
Date of Death	**DOD**	Date of death of this animal.
#Offspring		No. of immediate descendants of this animal (i.e., the number of other animals for which this animal is either the sire or dam).
#Males		No. of immediate descendants of this animal that are male.
#Females		No. of immediate descendants of this animal that are female.
#Unknown sex		No. of immediate descendants of this animal whose sex is not known.
#Families		No. of family records in which this animal is a parent.
#Litters		No. of litter records in which this animal is a parent.
#Full Siblings		No. of other animals with the same sire and dam as this animal. If either sire or dam is not known this is always 0.
#Litter Siblings		No. of other animals with the same sire and dam and date of birth as this animal. If either sire or dam is not known this is always 0.
#Half Siblings		No. of other animals that have either the same sire, or the same dam, as this animal, but not both. If either sire or dam is not known this is always 0.
#Paternal Siblings		No. of other animals with the same sire as this animal. If the sire is not known this is always 0.
#Maternal Siblings		No. of other animals with the same dam as this animal. If the dam is not known this is always 0.
#Siblings		No. of other animals with either the same sire, or the same dam, as this animal, or both.

Generation Number (1 + Parental Max)	**GN**	The generation number of this animal, defined as 1 plus the maximum of the sire and dam generation numbers. Founders always have GN 0.
Generation Number (1 + Parental Avg)	**GN**	The generation number of this animal, defined as 1 plus the average of the sire and dam generation numbers. Founders always have GN 0.
Age at Death		The time period between the date of birth and date of death of this animal.
Age If Living*		The current age of this animal, but only if it is considered as being alive (see 'Living' column).
Age At First Offspring		The time period between this animal's date of birth and that of its first known offspring.
Age At Last Offspring		The time period between this animal's date of birth and that of its last known offspring.
Generational Age		The average of the time period between this animal's date of birth and the dates of birth of its parents. Requires that 3 dates of birth to be known.
Age Of Sire At Birth		The time period between this animal's date of birth and that of its sire.
Age Of Dam At Birth		The time period between this animal's date of birth and that of its dam.

* ***Living*** and ***Age If Living***: These columns primarily exist to provide a way for large data sets to be sorted so as to bring all animals with a given current age, or all animals known to be living, together in the main record table. This makes it easier to tag the 'current' population (i.e., those that are alive), using a range-tagging operation. You may well want to do this if you are using PedScope to make mating recommendations.

Columns Computed By Pedigree Analysis

These are columns whose values can be computed by analyzing the pedigree of each animal at the same time as the inbreeding coefficient is computed.

The depth of ancestry (i.e., the number of generations) used in this analysis is a configurable parameter: the **Inbreeding #generations** setting on the **General** tab of the **Document Settings** window (**File|Document Settings**). You can either set a specific depth limit, such as 8 generations, or you can instruct PedScope to include all available ancestry by setting it to zero.

It may seem surprising but if you have a very large data set with many generations of ancestors present, PedScope will normally calculate these metrics ***much*** more quickly if you direct it to include all known ancestry. This is because of algorithmic optimizations that are used when analyzing the entire pedigree in a single step (rather than having to identify a separate sub-pedigree for each animal).

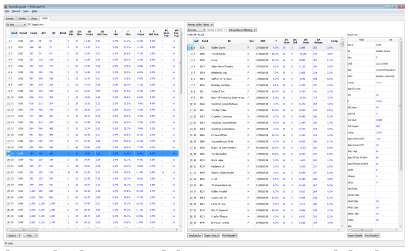
Columns	Symbol (if any)	Description
Inbreeding Coefficient	**F**	Wright's inbreeding coefficient for this animal, computed to the chosen depth of ancestry (see above).
Ancestral Inbreeding Coefficient		The ancestral inbreeding coefficient computed using the direct method of Ballou. This column is optional. To disable it choose **File|Document Settings**, click onto the **Compute** tab, then uncheck the **Compute ancestral inbreeding coefficients** option.
#Generations Seen **#Generations Full** **#Ancestors** **#Ancestors Unique** **Completeness**	**#G Seen** **#G Full** **#A** **#A Unique**	These are various metrics about the ancestors found when analyzing the pedigree of this animal. These are optional - to disable them choose **File|Document Settings**, click onto the **Compute** tab, then uncheck the **Compute individual pedigree metrics** option. **#Generations Seen** is the maximum depth of ancestry (#generations) seen. **#Generations Full** is the number of generations of ancestors that were full (no holes). **#Ancestors** is the total number of occurrences of ancestors seen. Ancestors occurring multiple times will add 1 to this value for every occurrence. **#Ancestors Unique** is the number of distinct ancestors seen. **Completeness** is the proportion of the pedigree that was full (no holes).

Statistics

The ***Stats*** tab of the main window calculates statistics after grouping the animals by date of birth. E.g. this allows you to compute statistics such as the average inbreeding coefficient on a year by year basis, of the number of animals born per decade.

Use of this feature requires that your source data includes the date/year of birth of each animal.

It should also be noted that this feature can use a large amount of memory, particularly if your source data contains a very large number of records. If you do not need this feature you can disable it - use the **Compute** tab of the **Document Settings** window (**File|Document Settings**).



Main Window - With stats pane with details of individuals falling within specific period

Above the main record table there is a pulldown menu where you specify the date reporting period - you can choose years, decades, months etc. The **Period** column shows the date range in which animals were born; the **Count** column holds the number falling in each period.

Above the main record table you will also find an option to restrict the statistics so that only tagged animals are included.

The Report Pane contains several different reports when displaying date period stats:

- ***Animals Within Period*** - lists all the animals born within the selected date period.
- ***Detailed Stats*** - shows detailed summary stats for all animals falling within the selected date period.
- ***Date Plot*** - lets you plot any one or two numeric columns derived from the animal records against time. E.g. to plot the average inbreeding coefficient over time.
- ***Correlation Chart*** - plots a scattergram for any two columns from the animal records for all animals falling within the selected date period. This operates in a very similar way to the correlation charts available for animal records.
- ***Frequency Chart*** - plots a frequency bar chart for any single column from the animal records for all animals falling within the selected date period. This operates in a very similar way to the frequency charts available for animal records.

Customization

PedScope provides various configuration/customization options. These are divided into two groups, accessed through separate windows.

The **Preferences** window is used for settings that are independent of any particular data file. It is accessed through the **File** menu. Changes to the preferences do not cause data to be (re)calculated. They have global effect - preferences are not specific to any particular data file. They tend to be related to the appearance and presentation of data, not the data themselves.

The **Document Settings** window is used for all settings that can alter the current data set in some way. It is accessed through the **File** menu. These settings affect what data is stored, or how it is stored, how it is calculated, or the accuracy with which it is calculated. These settings can be saved along with the imported pedigree information and computed data (see: Saving Data - Saving As A PedScope Data File).

Setting	How Set
Inbreeding depth (#generations)	Use the **General** tab of the **Document Settings** window.
Coefficient Display	Inbreeding and kinship coefficients can be displayed in three ways: as percentages (e.g. **12.5%**), or as numbers (e.g. **0.125**), or as fractions (e.g. **1/8**). To adjust this use the **General** tab of the **Preferences** window.
Coefficient Precision	Use the **General** tab of the **Preferences** window to adjust the precision (no. of decimal places) used for inbreeding and kinship coefficients.
Computation of Pedigree Metrics	Use the **Compute** tab of the **Document Settings** window to adjust the **Compute individual pedigree metrics** setting. This enables main table columns such as **#G Seen** and **#Ancestors Unique**.
Computation of Ancestral Inbreeding Coefficients	The **Compute** tab of the **Document Settings** window has an option to enable the computation of ancestral inbreeding coefficients.
Date Display	The display date format is set on the **System** tab of the **Preferences** window.
Statistics	You can enable/disable the date specific statistics tab using a setting on the **Stats Columns** tab of the **Document Settings** window. There are also options on the **Compute** tab to control which of the various statistics (mean, mode, median, etc.) are reported.
Main Table Columns	Use the **Main Columns** tab of the **Document Settings** window to select which columns to show in the main table. Note that the available columns are also affected by settings on the **Compute** tab.
Mate Recommendations Report	Use the **Mating Values** and **KV Weights** tabs of the **Document Settings** window to adjust parameters that affect the mate recommendations report.
ToolTips	There is an overall enable/disable setting for the tooltips that can appear over the main table and relations report: the **Show tooltips** setting on the **General** tab of the **Preferences** window. You can also adjust the information shown using the options on the **ToolTips** tab of the **Document Settings** window.
Fonts	Adjust the screen and printer font using the **General** tab of the **Preferences** window.
Graph Colors	Use the **General** tab of the **Preferences** window to adjust the colors used in graph and chart reports.
Row & Cell Highlighting	Use the **Highlighting** tab of the **Preferences** window.
Labels	The **General** tab of the **Preferences** window has options to control the usage of standard genetics symbols (e.g. **F** for the inbreeding coefficient) in column labels, and whether to use concise or longer labels.
Memory Usage	The **Large buffer size** and **Matrix representation** settings on the **General** tab of the **Document Settings** window affect the memory limits used in individual reports. The **Results cache size** setting on the **System** tab of the **Preferences** window controls the amount of memory that PedScope will use to store previously computed results. Increasing this value can speed up the program when switching to and fro between different reports - at the expense of using additional memory.
Speed	The **#Worker threads** setting on the **System** tab of the **Preferences** window controls the number of CPU 'threads' of execution that PedScope will attempt to use. Where possible the program uses multi-threading (parallel computations) to speed things up. The number of threads used is normally set so as to retain a responsive user interface. Increasing the **Results cache size** setting can speed up PedScope when switching between different reports.
Gene Drop Analysis	Use the **General** tab of the **Document Settings** window to set the #rounds to use when performing gene drop analysis.
Genome Uniqueness	Use the **General** tab of the **Document Settings** window to set the genome uniqueness threshold. For further information see Reports - Founder Metrics.
Defaults for Living Status column	Adjust the **Assume dead at age** setting on the **General** tab of the **Document Settings** window.
Use of	Use the checkboxes on the **General** tab of the **Document Settings** window. For further

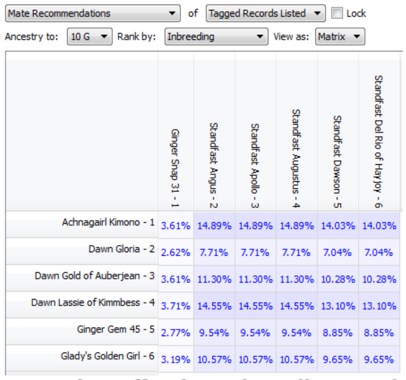
Family/Litter Records	information see Main Window - Other Record Types.
Sex-specific Offspring Counts	Use the **Compute** tab of the **Document Settings** window to control whether or not to have separate columns for the numbers of male/female offspring in the main table - use the **Compute offspring gender-specific metrics** setting.
Separated Maternal/Paternal Metrics	Some columns in the main table and relations report, such as the #sire/dam siblings and %blood can be tabulated separately for each parental 'side' of the animal's pedigree. To enable this, check the **Separated maternal/paternal metrics** setting in the **Compute** tab of the **Document Settings** window.
Generation Numbers	Use either of the settings in the **Compute** tab of the **Document Settings** window to enable the GN column. There are two settings, offering different ways to compute the generation number.
Relations Report Columns	Use the **Relation Columns** tab of the **Document Settings** window. Note that the available columns for the relations report are also affected by settings on the **Compute** tab.
Relation Descriptions	Use the **Precise relation descriptions** setting on the **General** tab of the **Preferences** window to adjust how genealogical relation descriptions are formatted in the relations report.
Number Precision	Use the **General** tab of the **Preferences** window to adjust the no. of decimal places used for numeric values (other than inbreeding/kinship coefficients - they are configured separately).
Terminology	PedScope uses a number of terms (words) such as 'Sex', 'Male', etc. To customize these terms, select **Preferences** from the **File** menu, then click on the **Terminology** tab. A number of preset choices are available for English-speaking breeders of cats, dogs, etc. Alternatively, select **Custom** from the dropdown menu then can change each term individually.

Reports

PedScope can display a wide variety of reports. Reports are displayed on the right hand side of the Main Window.

The following reports are available:

- Relations
- Founder Metrics
- Influential Ancestors
- Unknown Founders
- Gene Diversity Metrics
- Mate Recommendations
- Pedigree
- Correlation Chart
- Frequency Chart
- Relationship Matrices
- Families/Litters
- Record Details



	Ginger Snap 31 - 1	Standfast Angus - 2	Standfast Apollo - 3	Standfast Augustus - 4	Standfast Dawson - 5	Standfast Del Rio of Harjoy - 6
Achnagairl Kimono - 1	3.61%	14.89%	14.89%	14.89%	14.03%	14.03%
Dawn Gloria - 2	2.62%	7.71%	7.71%	7.71%	7.04%	7.04%
Dawn Gold of Auberjean - 3	3.61%	11.30%	11.30%	11.30%	10.28%	10.28%
Dawn Lassie of Kimmbess - 4	3.71%	14.55%	14.55%	14.55%	13.10%	13.10%
Ginger Gem 45 - 5	2.77%	9.54%	9.54%	9.54%	8.85%	8.85%
Glady's Golden Girl - 6	3.19%	10.57%	10.57%	10.57%	9.65%	9.65%

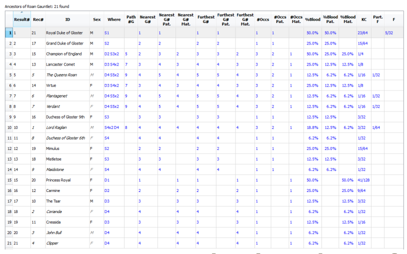
Prospective Offspring Inbreeding Matrix - Showing the inbreeding for every sire/dam combination within a chosen group of males and females

Reports - Relations

To view this report, choose ***Relations*** from the report selector (at the top of the report pane) in the main window.

Beneath the report selector are other options to choose the type of relations you want listed, and to restrict the results in various ways e.g. to limit ancestors to the maternal/paternal side, or to limit the depth of ancestry searched, etc. The matching relations are then listed. Some of the columns shown will be the same as those in the main record table. Other columns specific to the relations results are given below.

When you tabulate the relations for more than one animal at a time - e.g. list all the ancestors for a group of tagged records - the results will contain additional columns containing the min/max values of various of the above statistics, whilst other columns are no longer displayed as they are not applicable. E.g. the Kinship Coefficient for ancestors is not displayed, because it will often differ from one source record to another. But the min/max will still be shown.



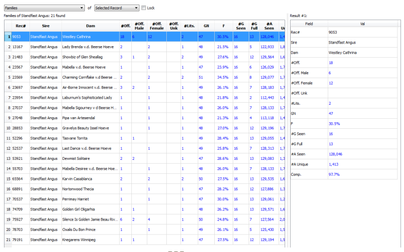
Ancestors Report - Showing columns for kinship and partial inbreeding coefficients with respect to founders

Column	Description
Where	For Ancestors, this gives the generation numbers and the side of the pedigree where the ancestor was seen. E.g. "S5x2 D7" means "twice on the sire side in generation 5, and once on the dam side in generation 7". When listing descendants this just gives the generation number (1=children, 2=grandchildren etc).
Path #Generations	When listing ancestors, if an ancestor is a common ancestor (only) this will give the minimum length of path to the common ancestor on one side of the pedigree and back down the other.
Nearest Generation#	When listing ancestors or descendants, this gives the closest generation at which the ancestor/descendant was seen.
Nearest Generation# Paternal	When listing ancestors, this gives the closest generation on the paternal side of the pedigree at which the ancestor was seen.
Nearest Generation# Maternal	When listing ancestors, this gives the closest generation on the maternal side of the pedigree at which the ancestor was seen.
Furthest Generation#	When listing ancestors or descendants, this gives the furthest generation at which the ancestor/descendant was seen.
Furthest Generation# Paternal	When listing ancestors, this gives the furthest generation on the paternal side of the pedigree at which the ancestor was seen.
Furthest Generation# Maternal	When listing ancestors, this gives the furthest generation on the maternal side of the pedigree at which the ancestor was seen.
#Occurrences	When listing ancestors or descendants, this is the number of times the ancestor/descendant was seen. This is >1 when there are multiple paths between the subject animal and the particular ancestor/descendant.
#Paternal Occurrences	When listing ancestors, the number of times this ancestor was seen on the paternal side of the pedigree.
#Maternal Occurrences	When listing ancestors, the number of times this ancestor was seen on the maternal side of the pedigree.
%Blood	When listing ancestors, this is the %blood contributed by that ancestor to the subject animal. When listing descendants, this is the %blood contributed by the subject animal to the descendant.
%Blood Paternal	When listing ancestors, this is the %blood contributed by that ancestor to the subject animal, but only when considering the paternal side of the pedigree.
%Blood Maternal	When listing ancestors, this is the %blood contributed by that ancestor to the subject animal, but only when considering the maternal side of the pedigree.
Kinship Coefficient (KC)	When listing ancestors, this is the kinship coefficient between the ancestor and the subject animal.
Partial Inbreeding Coefficient (Partial F)	When listing ancestors, if the ancestor is a founder, this is the partial inbreeding coefficient for that ancestor (i.e. that part of the inbreeding coefficient of the subject animal that is due to that founder ancestor).
Other Parent	When listing siblings, this is the other parent of a sibling (if not a full sibling).
Relation	When listing All Relatives, or Uncles/Aunts, or Nephews/Nieces, or Cousins, this is the genealogical description of the relation (e.g. 'Cousin Twice Removed', 'Half Brother' etc). You can sort the relatives by closeness of relation by clicking this column's heading. You can adjust these descriptions using the **Precise relation descriptions** setting on the **General** tab of the **Preferences** window.

Reports - Families/Litters

To view these reports, choose ***Families*** or ***Litters*** from the report selector (at the top of the report pane) in the main window.

Matching families/litters are listed in the report pane. Statistics on the numbers of offspring etc. are shown, along with other columns from the main table common to all offspring of each family/litter, e.g. the inbreeding coefficient.



Families Report

Reports - Record Details

To view this report, choose ***Record Details*** from the report selector (at the top of the report pane) in the main window.

Record details can be shown for both single and multiple animal records.

For single records, the record details is simply a list of all data stored for that record.

For multiple records (shown), the record details becomes a table of statistics for each field in the records, e.g. giving the min/max inbreeding coefficient etc. within the chosen records.

Field	Val	Min	Max	Mean	Std. Dev.
Age If Living					
Gen. # (1 + Par. Max)		46	52	50.20	2.27
COI		23.0%	26.7%	24.4%	1.3%
#G Seen		16	16	16.00	
#G Full		12	14	12.60	0.66
#A Seen		125,809	130,899	129713.60	1960.48
#A Unique		1,357	1,669	1576.30	125.67
Comp.		96.0%	99.9%	99.0%	1.5%
Age At First Off.		1y	3y	2y	<1y

Record Details Report

Reports - Correlation Chart

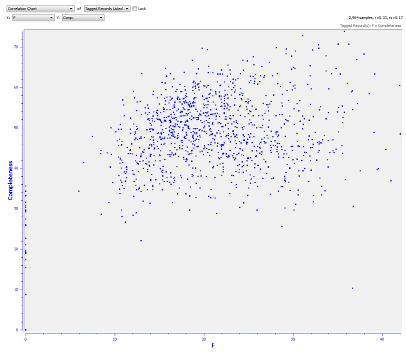
To view this report, choose ***Correlation Chart*** from the report selector (at the top of the report pane) in the main window.

A correlation chart (scattergram) is used to plot the correlation between any two column values over a set of animal records. Typically you would use this to chart the relationship between two metrics over a large number of tagged records, or perhaps over all records listed in the main record table.

You select which two columns from the main record table are to be plotted. One is plotted against the X axis, the other the Y axis. A dot is placed at the appropriate position on the chart for every source animal record that has values present in both columns.

In addition the program will compute the **Product Moment Correlation Coefficient** (**r**) and **Spearman Rank Correlation Coefficient** (**rs**) for the charted data. These are displayed above the chart.

A typical use of this feature would be to plot a correlation of, say, the inbreeding coefficient against a sample metric such as weight or height.



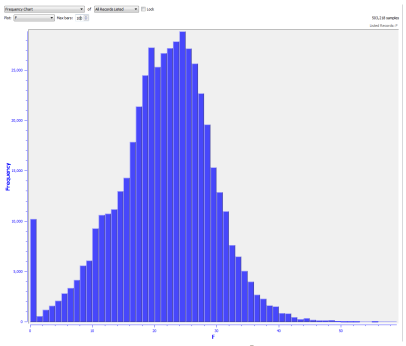
Correlation Chart

Reports - Frequency Chart

To view this report, choose ***Frequency Chart*** from the report selector (at the top of the report pane) in the main window.

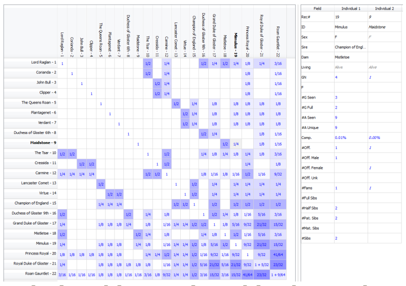
A frequency chart is a bar chart where each bar gives the number (frequency) of animals where a given metric (column) falls within a given range of values. This lets you visualize the frequency distribution of, say, inbreeding coefficients, within any group of animals.

You select which column from the main record table is to be charted, and the maximum number of bars to be plotted.



Frequency Chart

Reports - Relationship Matrices

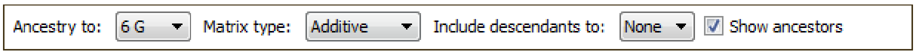


Relationship Matrix - This shows the A matrix for the stud "Roan Gauntlet" from Wrights original 1922 paper on inbreeding, with the "A" matrix values are shown as fractions

To view this report, choose ***Relationship Matrices*** from the report selector (at the top of the report pane) in the main window.

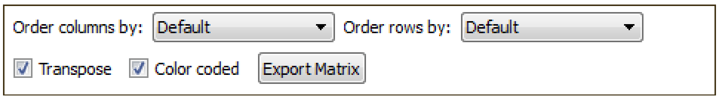
PedScope will compute and display various relationship matrices, such as the 'NRM' or 'A' matrix of additive genetic relations, between any group of related animals. The matrix is calculated for all of the source animal records, plus a given depth of their ancestors and, optionally, their descendants. The matrix can be saved (i.e. exported) including to spreadsheets.

When the report pane is set to display relationship matrices, several extra pulldown menus are available below the report selector:



- **Ancestry to** - here you select the #generations of ancestors of the source animals to be included when calculating the relations.
- **Matrix type** - select between **Additive**, **Dominance** and **Genomic**.
- **Include descendants to** - here you select the #generations of descendants of the source animals to be included.
- **Show ancestors** - this checkbox controls whether the ancestors used in the calculations are to be displayed in the visible matrix. This can be useful when you are interested in the relations between a small number of ancestors, but you don't want the screen cluttered up with a huge matrix showing hundreds or even thousands of their ancestors that have been used in calculating the relationships. Note that turning this option off has no effect on the memory requirements of the calculation (see below).

Below the displayed matrix are some further controls:



- **Order columns by** and **Order rows by** - these let you select the ordering of the rows and columns. The default order is the ordering used internally when computing the matrix, which is an ordering chosen to guarantee that all parents come before (i.e. left/above) their offspring. However you might find it easier to work with the displayed result if the rows/columns are sorted by, e.g., name. Note that changing the ordering will destroy any natural matrix structure e.g. with the additive matrix the diagonal always contains values that are 1 + the inbreeding coefficient. Changing the sort order will obviously mean that no longer holds. Changing the sort order does not recompute the matrix.
- **Transpose** - this will swap the rows and columns.
- **Color coded** - when this option is turned on, cells will be shaded according to their value. The higher the degree of relationship, the darker the shading. This provides a very quick way to see at a glance where the closest relationships are.
- **Export Matrix** - click this button to export the matrix e.g. to a PDF, text or XML file. The XML files generated by PedScope can be directly loaded into Excel and provide a very easy way to transfer a computed matrix.

When showing a relationship matrix the property pane to the right will display the details of the selected cell. Just click a matrix cell to have the details of the row + column animals displayed.

There are various other ways you can customize the display of matrices:

- On the **General** tab of the **Preferences** window you can specify whether coefficients are displayed as numbers, percentages or even fractions. For numbers and percentages you can set the number of decimal places shown.
- On the **General** tab of the **Document Settings** window you can specify the data storage type used internally for the matrix. This can affect both the memory required internally to store the matrix (which can be very large), and the speed of calculation.

Memory Usage: It is important to be aware that computing relationship matrices can require a **very large, potentially huge, amount of memory**. A relationship matrix contains as many rows and columns as the number of animals, and their ancestors and descendants that you have chosen to display. E.g. lets say you choose to display the additive matrix for 10 particular animals, plus 16 generations of their ancestors. And that once those ancestors have been found, this results in a total of 10,000 animals. A matrix with 10,000 rows and columns contains 100,000,000 cells. If each cell takes, say 8 bytes of storage (typical for a floating point number on many systems), this would be 800 Mb memory. Which is a lot. Whilst the program can (and does) use optimizations where it can to reduce the memory footprint (e.g. relationship matrices are symmetric, which at a stroke halves the memory requirement), it is still very easy to run out of memory. You can see the amount of memory used in the stats line displayed above the property pane when you have a relationship matrix visible. E.g. "N=1,696, memory=6.1Mb" says that the relationship matrix contains 1,696 rows + columns, and occupied 6.1Mb main memory. The program will not crash if you try to use more memory than is available, but it might become very slow or you might see an error message. Note also that if you compute the genomic relationship matrix, this will occupy 4 times as much memory for the same number of animals, because it has 2 rows and 2 columns for each animal.

Reports - Founder Metrics

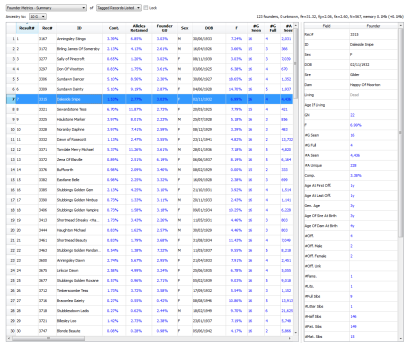
PedScope can analyse the founders of a single animal, or a group of animals, and list various measures of genetic diversity associated with them. Typically this would be used to analyse the founders of a group of animals that make up a 'current population' or a group from which future breeding is to occur. In the description that follows, 'current population' means either the currently selected animal, or the group of tagged animals, whichever has been chosen for the report.

There are 3 separate founder metrics reports available to choose from the report selector (at the top of the report pane):

- ***Summary***
- ***Contribution Detail***
- and ***Genome Uniqueness Detail***

The ***Founder Metrics - Summary*** report lists the founders and for each founder it gives its:

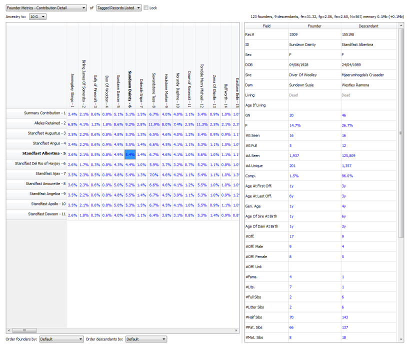
- **Contribution** - The proportion of the genes in the current population due to that founder.
- **Alleles Retained** - The proportion of the genes of that founder that have probably been retained in the current population.
- **Founder Genome Uniqueness** - The probability that genes from that founder are only present in, normally, 1 member of the current population. We say 'normally', because you can adjust the genome uniqueness threshold on the **General** tab of the **Document Settings** window.



Founder Metrics - Contribution Detail Report - listing the founders of a group of animals, showing the %contribution to the group, the %founder alleles retained, and the likelihood that each founders genes have only been passed down to at most one member of the group

As with all other tables in PedScope, you can sort the founders by clicking on the columns. It is therefore very easy indeed to work out which founders have had the main influence on any group of animals, and which have the least.

The other two founder metrics reports, ***Contribution Detail*** and ***Genome Uniqueness Detail***, display matrices that give the contribution and genome uniqueness values between every member of the current population and every founder. The matrix columns hold the founders, and the rows the members of the current population. The top two matrix rows always contain the summary value (i.e. the contribution or genome uniqueness with respect to the whole group), and the alleles of each founder retained by the whole group - the same values as given in the summary table. The properties pane comprises two columns and contains the record details for the currently selected matrix cell.



Founder Metrics Contribution Detail Report - Matrix with founders in columns and group members in rows, giving the %contribution of each founder to each descendant

All founder reports let you choose the depth of ancestry to be considered when building the report. PedScope searches the pedigrees of all animals in the current population to that depth to provide a list of ancestors, and then locates animals within that group that are founders ***with respect to the group***. This is an important distinction; see Population Genetics - Founders.

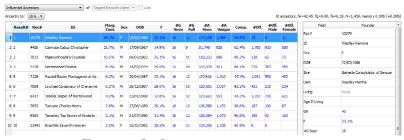
All founder reports also include the following measures of genetic diversity, which are single values computed for the whole of the current population:

- **fe** - the number of founder equivalents
- **fg** - the number of founder genome equivalents
- **fa** - the number of effective ancestors
- **N** - the total number of animals analysed for the report (current population + ancestors)

You can customize the definition of genome uniqueness. By default, a gene from a given founder is considered to be inherited uniquely when it is present in only one member of the current population. You can increase this threshold, e.g. you could choose to have PedScope treat the presence of the gene in at most 3 animals as meaning 'unique'.

The probability of alleles retained, and genome uniqueness, are computed using a gene drop analysis. You can customize the number of simulations used to compute these values.

Reports - Influential Ancestors



Influential Ancestors Report

To view this report, choose ***Influential Ancestors*** from the report selector (at the top of the report pane) in the main window.

This report displays the most influential ancestors of a group of animals. To use this report you must first tag the animals whose influential ancestors you want to locate. This report does not operate with a single ancestor.

The 'influential ancestors' are those ancestors with the greatest contributions to the computation of the number of effective ancestors (**fa**). The table displayed includes a column that gives the marginal contribution of each ancestor to the current population.

The algorithm for determining influential ancestors [1] starts with the ancestor with the highest genetic contribution to the current population, and then proceeds to list other ancestors in decreasing order of their genetic contribution, whilst attempting to ensure that contributions already explained by influential ancestors previously found do not get considered twice.

The number of effective ancestors (**fa**) is also computed and displayed above the report along with the numbers of founder equivalents and founder genome equivalents.

For further information about the number of effective ancestors and related metrics, see Population Genetics.

---

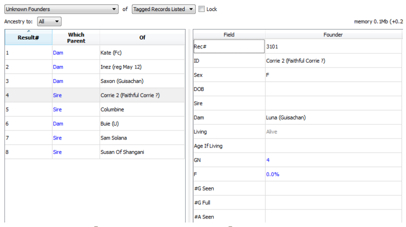
1. Boichard, D, Maignel L, and Verrier, E. The value of using probabilities of gene origin to measure genetic variability in a population. Genetics Selection Evolution, 1997; 29(1): 5-23.

Reports - Unknown Founders

To view this report, choose ***Unknown Founders*** from the report selector (at the top of the report pane) in the main window.

An 'unknown founder' in PedScope refers to the situation when one parent of an animal is not known, but the other is. The unknown parent is considered to be an unknown founder.

This report will list the unknown founders of the current population, within a given depth of ancestry.



Unknown Founders Report

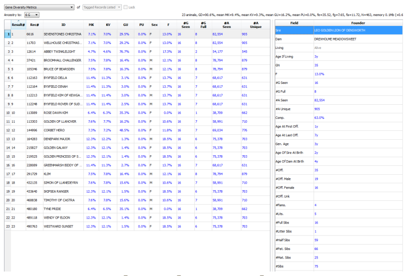
Reports - Gene Diversity Metrics

To view this report, choose ***Gene Diversity Metrics*** from the report selector (at the top of the report pane) in the main window.

PedScope can tabulate various gene diversity metrics of a group of animals as used in the separate mate recommendations report.

Typically this would be used to study a group of animals that make up a 'current population' or otherwise a group from which future breeding is to occur.

In the description that follows, 'current population' means the group of tagged animals from the main record table.



Gene Diversity Metrics Report

Column	Symbol	Description
Mean Kinship	**MK**	The mean kinship of an animal is the mean of its kinship coefficients with every animal in the current population, including itself. The lower the MK value, the less an animal is related to other members of the current population, 'on average'. Informally, it makes sense to breed from such animals in order to preserve genetic diversity. Note that two animals with low MK may nevertheless be closely related themselves, so selecting matings on this basis alone can still give rise to highly inbred offspring.
Kinship Value	**KV**	Kinship value is simply a weighted variant of the mean kinship, weighted according to the age structure of the population. For this to work, the program needs additional information: (a) the date of birth of each animal (from which its age is determined), and (b) the reproductive values of different aged animals, as relevant to the particular species (so-called 'Vx' values) [see Fisher, R. The Genetical Theory of Natural Selection, 1930]. Clearly, an animal in the current population that is no longer reproductive e.g. a female nearing the end of her life, is of no value in maintaining future genetic diversity, and yet calculations based on MK do not take this into account. Likewise it follows that if you have two animals with the same, low, mean kinship, but one is older than the other, then a greater priority might be given to breeding from the older animal first, to lessen the risk that its genes are lost for good. Using KV in such mating decisions allows these factors to be incorporated into the mate ranking algorithm, whereas ranking by MK makes no distinction. KV can be used in the mate recommendations report. You can enter Vx data using the **Document Settings** window - see Vx Data.
Genome Uniqueness	**GU**	An animal's genome uniqueness is the likelihood that it contains founder alleles that are not present in any other animal (normally) in the current population. Informally, if an animal is the only one carrying a particular founder's genes, then you'd likely want to breed from that animal so as not to lose those genes from the population forever. We say 'normally' here because you can actually configure what is meant by genome uniqueness, in that you can define that number of animals in the current population for which, if any given founder's alleles are only present in that number, or fewer, of the current population, then it is considered 'unique'. A better term might be 'genome rarity' or 'founder rarity' but the term 'genome uniqueness' is already established in the literature. Genome uniqueness is determined using a gene drop analysis.
Pedigree Unknown	**PU**	The 'pedigree unknown' value for an animal is that proportion of its founders that are unknown. This relates only to situations where an ancestor has one parent missing in the ancestry; such missing ancestors are considered to be unknown founders, missing from the pedigree. Sometimes it can be desirable to avoid breeding from animals with a large degree of unknown pedigree, for this reason the PU value can be used in the composite mate ranking algorithm.

Reports - Mate Recommendations

To view this report, choose ***Mate Recommendations*** from the report selector (at the top of the report pane) in the main window.

PedScope makes breeding recommendations by studying potential pairings and ordering them according to a single value reflecting the respective merit of the pairing. In this way it can be used as a tool to guide population management decisions. It provides various ways of calculating the merit for each pairing.

As with the Gene Diversity Metrics report, to use this feature you first need to define a group of animals that represents the available sires and dams. You do this by tagging the sires and dams in the main record table.

If you want to include all the living members of the population it can help if you first sort the main record table by the **Age if Living** or **Living** columns, if available. This will bring all the living animals together so you can tag them with a range-tagging operation.

Next, you choose the depth of ancestry to be considered in the calculation using the **Ancestry to** pulldown. This includes the option of using all known ancestry.

Finally, you choose the method by which the pairings are to be ordered using **Rank by** pulldown. The following table describes the different rankings in detail, and gives some of advantages and disadvantages of each.

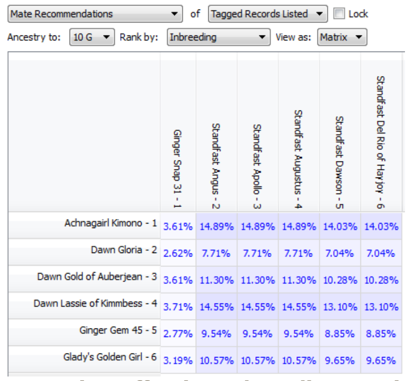
<table>
<tr><th>Ranking Method</th><th>Description</th></tr>
<tr><td>Inbreeding</td><td>
Breeders of fancy animals such as dogs and cats are often particularly interested in limiting the inbreeding of the offspring of a prospective mating.

To help with this one of the available ranking methods is simply the coefficient of inbreeding of the potential offspring.

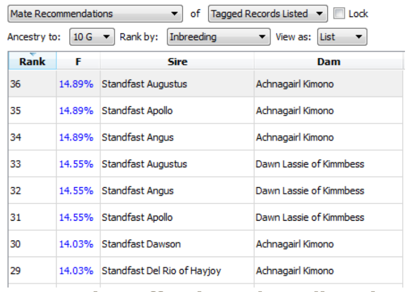
The results can be listed either in a table, ordered by ranking, or as a color-coded matrix of inbreeding coefficients with the sires in columns and the dams in rows.

If your aim in using the mate recommendations report is to maximize genetic diversity over the longer term, simply minimizing the inbreeding in immediate breeding decisions is not necessarily the best way to go.

Better results will be obtained by using a ranking method that uses Kinship Value (KV), such as the Parental KV or Composite rankings (see below).



Prospective Offspring Inbreeding Matrix - Showing the inbreeding for every sire/dam combination within a chosen group of males and females



Prospective Offspring Inbreeding List - Ranking list of prospective pairings, ordered by offspring inbreeding
</td></tr>
<tr><td>Gene Diversity Delta (GD)</td><td>
Gene diversity (GD) is a measure applied to a whole population that expresses the diversity of genes present. Gene diversity Delta is the change in gene diversity of the population as a result of the addition of a single offspring from the prospective mating.

If every animal in a population is unrelated, then by definition you would have the maximum genetic diversity possible for a population of that size.

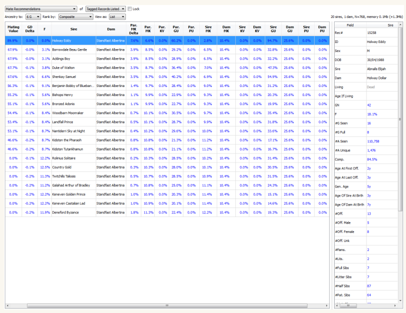
A common goal of population management programmes for threatened species is to maximize the genetic diversity of the available population.

When ranking by GD Delta what PedScope does is this: for each possible pairing, it computes the GD that would result by addition of a single animal with given parents to the current population, and then compares this to the GD of the 'old' (i.e. current) population. An increase in GD as a result of the pairing is a good thing, a decrease is a bad thing.
</td></tr>
<tr><td>Gene Value Delta (GV)</td><td>
This is the change in gene value of the population as a result of the addition of a single offspring from the prospective mating. GV is a function of the average kinship value of the population. GV is to KV what GD is to MK.

It should be preferred over GD when kinship values are available because it takes into account the age structure of the current population.

This method requires that you have 'Vx' data set up, in order to compute the kinship values. For further information see Vx Data.
</td></tr>
<tr><td>Parental MK</td><td>
This is the mean of the parental mean kinships. I.e. PedScope computes the mean kinship of both sire and dam (with respect to the current population) and then ranks the prospective matings according to the mean of these two values for each pairing.

A lower value is better (because, lower values prioritize mating of animals that share fewer of their genes with the rest of the population).
</td></tr>
<tr><td>Parental MK Delta</td><td>
This is the difference in the mean kinships of the sire and the dam.
</td></tr>
</table>

	It can be detrimental to pair animals with widely differing mean kinships because this has the effect of combining rare gene lines with over-represented lines. A lower MK delta is better, and ranks higher.
Parental Genome Uniqueness	This is the average of the genome uniqueness of the sire and the dam. It is the total probability, for all founders, that the animal is the only one (normally) in the current population containing genes from any one founder. Therefore the higher an animal's 'genome uniqueness', the greater the chance that if that animal is not bred from, then under-represented gene lines will disappear forever from the current population.
Parental KV	This is the mean of the parental kinship values. The kinship value of an animal is related to the MK, but is weighted according to the age structure of the current population. This can be a better metric upon which to base breeding decisions than the mean kinship, but requires that you also have reproductive values for the different age classes of the current population ('Vx data'), and - obviously - you have the age data of the current population. E.g. the current population may include animals that are past reproductive age. Clearly such animals can never be bred from, so it is better if the genetic diversity contributed by those animals to the current population be discounted when making breeding decisions. Use of KV values enables this. This method requires that you have 'Vx' data set up, in order to compute the kinship values. For further information see Vx Data.
Composite	This is where a weighted combination of any of the other breeding metrics above, alongside the offspring inbreeding coefficient, is used. This is probably the most ranking powerful method. It lets you combine the various measures in whichever way you like, weighted according to the values you wish to prioritize. You can also set a limit threshold for the offspring inbreeding, so as to remove from the rankings any pairings that would result in excessively highly inbred animals, irrespective of the other merits of the pairing. To customize the weightings used by the composite ranking, use the **KV Weights** tab of the **Document Settings** window. For further information see Population Genetics.  ***Mate Recommendations Report - Pairings ranked by composite breeding value***

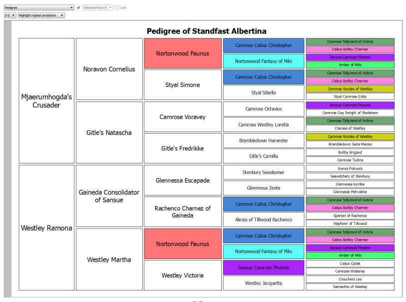
Reports - Pedigree

To view this report, choose ***Pedigree*** from the report selector (at the top of the report pane) in the main window.

You can display a simple pedigree chart for the currently selected animal in the main record table and include between 2 and 6 generations of ancestors.

Use the **Highlight** pulldown to select options for highlighting ancestors:

- **No highlight** - as it says.
- **Highlight repeat ancestors** - ancestors that occur more than once within the displayed generations are highlighted with a unique color.
- **Highlight common ancestors** - ancestors that occur on both sides of the pedigree within the displayed generations are highlighted with a unique color.



Pedigree

Population Genetics

This section describes some of the concepts of population genetics that are used by this program.

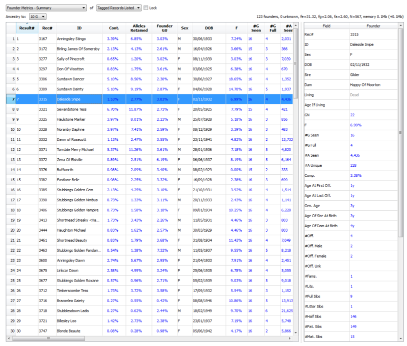
If you are knowledgable in this field, you probably don't need to read this section.

But if terms such as 'mean kinship' and 'founder equivalents' are unfamiliar to you, then you ***do*** need to this section, as understanding of these concepts is necessary in order to get the most out of a product such as PedScope.

Founders

A ***founder*** within a set of pedigree data is defined as an animal where neither the sire nor dam is known. Such an animal may truly be a 'founding ancestor' of a breed or 'population' in the sense that it is not related to any other founder, or it may be related - possibly closely - to other members of the group but the details of its lineage are not known. Within PedScope founders are treated as unrelated and non-inbred.

Many features of PedScope involve determining the ancestors of an animal or group, and then computing various metrics or reports based on them. In most situations PedScope gives you a way to artificially restrict the depth of ancestry that is considered. E.g. you may choose to limit inbreeding calculations to 16 generations. Whenever such an artificial cutoff is applied, the definition of 'founder' takes on a slightly different meaning: it means an animal, ***within the subset of the population that comprises the animal(s) under study and however many generations of their ancestors that has been selected***, that either has no known sire and/or dam or whose sire and/or dam, even though known in the entire population, do not fall within the subset under study. It is as if the program 'pretends' that the deeper ancestry is simply unknown.



Founder Metrics - Contribution Detail Report - listing the founders of a group of animals, showing the %contribution to the group, the %founder alleles retained, and the likelihood that each founders genes have only been passed down to at most one member of the group

Alleles, Genes & Identity by State/Descent

Most animals have two sets of chromosomes. They have one 'copy' of each gene on each chromosome. Such genes are said to be ***autosomal***. The particular location on a chromosome where a gene is found is called its ***locus***. An animal therefore has 2 'copies' of a gene for any particular locus. An ***allele*** is one of two or more forms of a gene. The presence of different combinations of alleles can affect physical traits in the animal, such as coat color in some animals. This description is a deliberate simplification, but sufficient for our purposes. E.g. we ignore sex chromosomes etc.

If both alleles for a given locus are the same, the animal is said to be ***homozygous*** for that gene. The phrase ***identical by state*** (IBS) means the same thing: the alleles are functionally the same. If the alleles are different, they are said to be ***heterozygous***.

A related term is ***identical by descent*** (IBD). Two alleles are said to be IBD if one of them is a physical copy of the other, or if they are both physical copies of the same allele in a particular ancestor. IBD implies IBS but not vice versa.

Kinship

The ***kinship coefficient*** (KC) (also known as ***coancestry***) between any two animals is the probability, for any particular locus, that an allele selected randomly from one animal is IBD to an allele selected randomly from the other.

An animal gets half its genes from its father, half from its mother. So that means the that 'kinship coefficient' between an animal and either of its parents must be one half (i.e. 50%, or 0.5), right? Wrong. The KC between an offspring and either parent is one quarter i.e. 25%. Why is this? First, there is a 50% chance that any allele chosen at random in an animal is from a particular parent (the sire, say). And secondly there is then a 50% chance that that is the same as any particular allele chosen randomly in the sire. So, there is a 0.5 * 0.5 chance i.e. 0.25 or 25% that they are IBD.

Inbreeding Coefficients & Common Ancestors

Inbreeding occurs when an animal has one or more ***common ancestors***.

The term 'common ancestor' has a very precise meaning in population genetics. It means an ancestor that is present at least once on both sides of an animal's pedigree. It does ***not*** mean an ancestor that just 'occurs a lot'.

When an animal has common ancestors, this raises the possibility that for any particular gene locus, both alleles present are in fact physical copies of the exact same allele from one of its common ancestors. The inbreeding coefficient is the value of this probability. Or, put another way, it is the probability that both alleles are IBD.

The inbreeding coefficient as in common use today was first defined in 1922 in a paper by Dr Sewall Wright [1]. For this reason it is often known as 'Wrights Inbreeding Coefficient'.

When an animal has no common ancestors, its inbreeding coeffficient is zero. The symbol '**F**' is commonly used in scientific literature to mean an animal's inbreeding coefficient.

The inbreeding coefficient can be computed by hand only for relatively simple pedigrees. For deep pedigrees it can only realistically be done by computer. Calculating inbreeding for very large pedigrees is a computationally intensive process and if not implemented very carefully can take a long time, if it is correct at all. The computation of inbreeding coefficients in PedScope is ***amazingly fast***. E.g. to compute all inbreeding coefficients to the maximum depth possible for the UK Golden Retriever pedigree - some 500,000+ dogs going back up to 65 generations - took just 7 seconds [2].

Partial Inbreeding Coefficients

A ***partial inbreeding coefficient*** of an animal with respect to a specific founder is the probability that the animal is IBD for an allele descended from the specific founder. It is that part of the animals inbreeding that is due to that founder.

PedScope can tabulate partial inbreeding coefficients when you list ancestors. For further information see Reports - Relations.

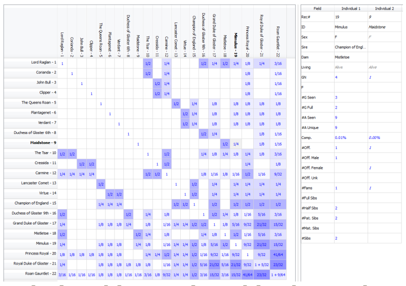
Ancestral Inbreeding Coefficients

The ***ancestral inbreeding coefficient*** of an animal is the cumulative proportion of an animal's genome that has been previously exposed to inbreeding in its ancestors [3]. Ancestral inbreeding coefficients can be tabulated in the main record table.

Relationship Matrix

A ***relationship matrix*** is used in population genetics to express the genetic relations within a set of animals and their ancestors. Such matrices are always square: they have one row and one column for each animal. It is accepted practice that the rows/columns are ordered such that offspring are always preceded by their parents. Thus the first row/column will always hold a founder of the group under study.

The most commonly used matrix is the matrix of ***additive genetic relationships***. This is sometimes referred to in population genetics literature as the 'A' matrix. Each cell in this matrix, say the intersection of row I and column J, gives the additive genetic relationship between animals I and J. When I is not equal to J, i.e. for off diagonal values, this is equal to twice the kinship coefficient. When I is equal to J, i.e. for the diagonal, this is equal to the inbreeding coefficient of animal I, plus one.



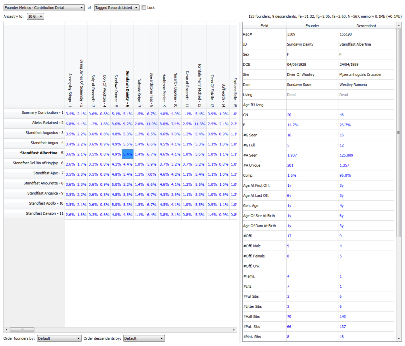
Relationship Matrix - This shows the A matrix for the stud "Roan Gauntlet" from Wrights original 1922 paper on inbreeding, with the "A" matrix values are shown as fractions

Founder Equivalents & Founder Genome Equivalents

The founders of a population under study contain all the genetic diversity available to be inherited by their descendants. However, not necessarily all of the genetic variation present in the founders may have 'made it' to the current population, due to genetic drift, selection and inbreeding.

Genetic drift refers to the loss, by chance, of genetic diversity present in the founders. E.g. when a founder has only one offspring, but that offspring is then the ancestor of one or more members of the current population, then by definition at least half of the genetic diversity that was present in that founder must have been lost. That is because its sole offspring only inherited half of it. Even when a founder has many offspring and from which there may be large numbers of descendants in the current population, it is not likely that all its genes will still be present.

Larger populations, with larger numbers of founders, large families and little or no selection i.e. truly random mating usually maintain genetic diversity. Genetic diversity can be increased through mutation and migration i.e. the introduction of truly unrelated 'new' founders. But within smaller populations, such as 'closed' populations of fancy animals or with endangered species where there are very limited numbers of animals remaining, the extent of retained genetic diversity can be limited, and the extent to which different founder's are represented within the current population can vary, sometimes greatly.



Founder Metrics Contribution Detail Report - Matrix with founders in columns and group members in rows, giving the %contribution of each founder to each descendant

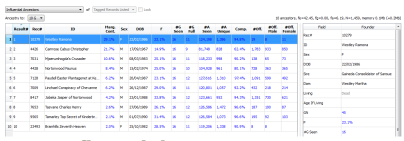
The number of ***founder equivalents*** (Lacy [4]) is a measure of the genetic diversity of a current population. It is the number of equally contributing founders that would be expected to produce the same level of genetic diversity as the current population. It has a standard symbol in population genetics literature: '**fe**'. In a large population with no selection and in the absence of genetic drift fe remains relatively constant because founder contributions do not vary much from one generation to another. When there is extensive selection fe can lose value, such as with fancy animals when popular sires can give rise to huge numbers of offspring and hence magnify the representation of their founders within the current population at the expense of those of unpopular sires.

The number of ***founder genome equivalents*** (Lacy [4]) is a related measure which, unlike the number of founder equivalents, takes genetic drift into account. It is the number of equally contributing founders with no random loss of alleles in the offspring that would be expected to produce the same level of genetic diversity as the current population. It has standard symbol '**fg**'. The calculation of fg is more involved than fe because it requires knowledge of the extent to which each founders alleles are present in the current population i.e. the degree of genetic drift. This can be calculated using a gene drop analysis. fg will always be less than or equal to fe.

Effective Ancestors

Another useful metric is the number of ***effective ancestors*** (Boichard et al [5]). It is similar to the number of founder equivalents except that it also takes into account bottlenecks in the pedigree.

It does this by working out the marginal contribution of each ancestor - not each founder - to the current population to find the ancestor with the most influence, and repeats the process taking care not to consider the contribution of any previously identified influential ancestor more than once.



Influential Ancestors Report

The result is a list of the most influential ancestors, and a metric, the number of effective ancestors. It has standard symbol in population genetics - '**fa**'. This number will always be less than or equal to **fe**.

Mean Kinship (MK)

The ***mean kinship*** (symbol: **MK**) of an animal within a group of animals is the mean of its kinship coefficient with every other member of the group, including itself. If an animal's MK is low, this means it is less related, on the whole, to the rest of the population than an animal with a higher MK.

MK plays an important part in breeding decisions in programmes designed to maintain genetic diversity in small populations. Leaving aside other factors, if you had to choose between one sire and another for a mating decision, it would be better to choose the one with lower MK. But, it is inferior to a related measure, kinship value.

Kinship Value (KV)

The ***kinship value*** (symbol: **KV**) is a weighted variant of mean kinship. The weighting used is the reproductive value of the animal concerned; it is age-specific and defined as the expected future lifetime reproduction (Fisher's reproductive value 'Vx' [6]).

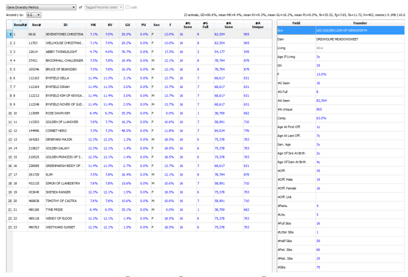
Clearly, an animal in the current population that is no longer reproductive e.g. a female nearing the end of her life, is of no value in maintaining future genetic diversity (because she can no longer produce offspring to add to the population), and yet calculations based on MK do not take this into account. Likewise it follows that if you have two animals with the same MK, but one is older than the other, then - and ignoring any other factors - a greater priority ought to be given to breeding from the older animal first, to lessen the risk that its genes are lost for good.

In order to compute KV, additional information is required. Firstly the current age of each animal is needed. Which means, in practice, that the input data to PedScope must include the date/year of birth. Secondly, the program needs the weights to be assigned to each animal. PedScope provides 2 ways in which the weights can be entered for KV calculation. For further information see Vx Data.

Gene Diversity (GD)

The ***gene diversity*** (symbol: **GD**) of a population is defined as 1 minus the 'mean MK' (i.e. the mean of the MK's of every animal in the current population, including with itself). The lower the mean MK of all animals in a population, the less related, on average, they are to each other. GD is simply a different way to view the same number.

Maximizing GD can be a good basis for a breeding programme e.g. for an endangered species with a small population but with known ancestry, though it is inferior to GV.



Gene Diversity Metrics Report

Gene Value (GV)

The ***gene value*** (symbol: **GV**) of a population is defined as 1 minus the 'mean KV' (i.e. the mean of the kinship values of every animal in the current population, including with itself). GV is to KV what GD is to MK. Since KV's are, in effect, an improved form of MK, so GV is an improvement upon GD.

If you have age-structure data for the current population along with reproductive values, GV should normally be preferred over GD as the method of ranking breeding decisions.

Genome Uniqueness

The ***genome uniqueness*** (GU) of an animal with respect to a current population of which it is part is the probability that it contains founder alleles not present in any other single (normally) animal in that current population. GU can also be computed with respect to a specific founder in which case it is the probability that the genes from that founder are inherited 'uniquely'. GU can be used as a factor when making breeding/mate recommendations.

Informally, if an animal is the only one carrying a particular founder's genes, then you'd likely want to breed from that animal so as not to lose those genes from the population forever.

We say 'normally' here because PedScope lets you configure what is meant by genome uniqueness, in that you can define that number of animals in the current population for which, if any given founder's alleles are only present in that number, or fewer, of the current population, then it is considered 'unique'. For further information see Customization. A better term might be 'genome rarity' or 'founder rarity' but the term 'genome uniqueness' is already established in the literature. Genome uniqueness is determined using gene drop analysis.

Gene Drop Analysis

A ***gene drop analysis*** (MacCluer et al [7]) is a computer simulation technique for analyzing a group of animals comprising a current population and their ancestors.

The founders of the group are first identified. The computer then simulates Mendelian inheritance by 'dropping' a gene, comprising pair of unique alleles for each founder, 'through' the pedigree. The actual alleles that the current descendants have 'received' are then counted up. The whole process is then repeated, usually many thousands of times. By counting the frequencies of the alleles that the descendants end up with, it is possible to work out good approximations for various metrics that otherwise would be difficult to work out exactly.

PedScope uses gene drop analysis for the computation of genome uniqueness, and the proportion of alleles retained in the current population from any given founder (which in turn is required for the calculation of the number of founder genome equivalents).

---

1. Wright, S. Coefficients of inbreeding and relationship. American Naturalist 56: 330�338, 1922.
2. Using Windows 7, 64 bit, AMD Phenom II X6 1035T processor 2.6Gz, 8Gb memory.
3. Ballou, JD. Ancestral Inbreeding Only Minimally Affects Inbreeding Depression in Mammalian Populations, Journal of Heredity 1997, 8:169-178.
4. Lacy, RC. Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. Zoo Biology 8:111-124, 1989.
5. Boichard, D, Maignel L, and Verrier, E. The value of using probabilities of gene origin to measure genetic variability in a population. Genetics Selection Evolution, 1997; 29(1): 5-23.
6. Fisher, R. The Genetical Theory of Natural Selection, 1930.
7. MacCluer, JW, VandeBerg JL, Read B and Ryder OA. Pedigree analysis by computer simulation. Zoo Biology 5, 147/160, 1986.

Vx Data

Kinship values, and the associated mating ranking metric 'Offspring GV Delta' (see Reports - Mate Recommendations) require reproductive values of the animals concerned. This data is age-specific and defined as the expected future lifetime reproduction (reproductive value 'Vx' [see Fisher, R. The Genetical Theory of Natural Selection, 1930.]).

There are 2 ways you can supply this data to the program:

- In a file that is automatically loaded when you import your data.

  When PedScope loads data from a text/CSV file, it always checks for the existence of a file with the same name as the import file but with '-kvweights' inserted in front of the import file extension. E.g. if you import a file called "mystudbook.txt", then PedScope will also look for a file "mystudbook-kvweights.txt".

- OR, entered separately through the document settings.

  Choose **File|Document Settings**, click on the **KV Weights** tab and key in the Vx data (or click **Load** to read from a file).

In both cases the data should be formatted as separate lines giving SEX,AGE,VXVALUE. E.g. 'm,3,0.8' would mean that males of age 3 years have a Vx value of 0.8.

A longer example:

```
m,3,1.0
m,4,0.8
m,5,0.4
m,6,0.35
m,7,0.3
m,8,0.3
m,10,0.1
m,12,0.0

f,2,1.0
f,3,0.9
f,4,0.8
f,5,0.5
f,6,0.3
f,7,0.1
f,8,0.0
```

Saving Data

PedScope provides several options for saving and exporting data and reports.

Saving As A PedScope Data File

You can save the entire state of a displayed pedigree data set, along with all the current Document Settings and computed data (inbreeding coefficients etc.) in a single file that can be easily reopened later. Opening such a saved data file is usually much quicker than reloading the original import file.

To save the data in this format, choose **File|Save** (or **File|Save As**).

To subsequently reload such a data file, choose **File|Open**, change the file type to 'PedScope data files', then locate the saved file.

Exporting

You can export individual reports, data tables etc. in a variety of file formats. To do this you just click the **Export** or **Save** button which can be found near the bottom of most the main window panes:

- ***Data tables and matrices*** - PDF, PostScript, text/CSV, XML (for direct import into spreadsheets such as Excel), HTML and Open Document Format (.odt).
- ***Graphs*** and ***Pedigree charts*** - PDF, PostScript, JPEG, PNG, Scaled Vector Graphics (.svg, for websites), and other bitmap image formats.

Printing

You can print the main record table, and most reports, by clicking the **Print** button (beneath the table/report).

Saving As PDF/Image Files

Instead of printing to a physical printer, you can also 'print' to a PDF or image file - by exporting to a PDF file.

Pedigree Margins

When printing the pedigree report, PedScope sets default paper margins depending on your printer and its setup.

In most cases, this should produce a sensible and consistent margin on all sides of the paper.

But if not, you can change this - click the margins button ().

Understanding Printer Margins

This section contains a more detailed explanation of how PedScope sets paper margins, and how margin settings work with your printer and chosen paper size.

Most printer devices have physical limitations on their printing capabilities. Obviously the supported paper sizes vary. But in addition, most printers are not physically capable of printing on the entire sheet of paper: there are usually parts of the sheet that the printer simply cannot physically print on. You can instruct PedScope to attempt to print onto that maximum part of the paper by selecting the **Full page** option within the **Print Margins** window.

Your computer will also have a default margin setting for the chosen printer. This should be a sensible default; its actual size is system specific. You can instruct PedScope to use the default margins by selecting the **Use printer default margins** option within the **Print Margins** window. This is the normal setting and an in ideal world will produce a good result without you having to change anything.

But you can opt to override the default margin selection by disabling both the **Full page** and **Use printer default margins** options in the **Print Margins** window, and then entering the actual margins you would like. Measurements can be entered in a variety of units including centimeters (e.g. you could enter `1.2cm`) and inches (e.g. `0.25in`). Note that setting margins smaller than the physical minimum margins that your printer is capable of generally will not produce acceptable results - you will end up with parts of the printout 'chopped' off. You may find you need to experiment with different sized margins until you get a pleasing result.

FAQs

Listed here is a selection of some of the more commonly asked questions about PedScope.

Additional questions and answers can be found on our web site.

Q: **How can I interactively enter data?**
A: You can't - sorry! This software is purely intended as a standalone program for analyzing a pedigree data set which you import from a text or CSV file. It cannot be used to enter data e.g. onto a database.

Typical usage of this software is for analyzing a data set stored using a separate database or software, from which you export the data into CSV/text format for import into this software.

If you would like to use this software in conjunction with other sources of data please contact us.

Q: **How do I customize PedScope?**
A: Customizations that are not specific to the current pedigree data file - display preferences etc. - are controlled using the **Preferences** window - select **Preferences** from the **File** menu.

Customization specific to the pedigree data - such as the user defined fields - are controlled using the **Document Settings** window - select **File|Document Settings**.

See Customization.

Q: **How do I uninstall PedScope?**
A: ***Windows:*** Open Control Panel and select 'Uninstall a Program' or 'Add/Remove Software'.
Mac: Just drag PedScope to the trash can.

Q: **What extra do I get with the full (paid) version?**
A:

- It is not limited to 30 days use
- It works with data files of up to 20,000,000 records whereas the free trial is limited to 500
- Export and print facilities are enabled in the paid version (they are disabled in the trial except when used with the bundled examples or simulated pedigrees)
- Access to support and updates

You can test any of the features of the full version even if you only have the free trial by opening one of several inbuilt pedigree data files (use **File|Examples**), or by creating a simulated pedigree (use **File|Simulate Pedigree**). These are not subject to the trial version record limit.

Q: **How do I place an order?**
A: See our web site.

Q: **My question isn't answered here, where else can I look?**
A: Please see the ***Frequently Asked Questions*** page on our web site, which we update regularly. This addresses many additional topics and questions to those listed here.

Q: **My question isn't answered in the FAQs, neither here nor on the PedScope web site. How can I contact product support?**
A: Please see ***Product Support*** on our web site.

Product Support

To contact customer service please see the ***Product Support*** page on our web site.

Before contacting customer service, be sure to check the ***FAQs*** on our web site first. This covers many commonly asked questions, and may well address the problem or query that you have.

Release Notes

System Requirements

PedScope requires Windows 8, Windows 7 or Windows Vista (either 32 or 64 bit).

Revision History

VERSION	RELEASE DATE	DETAILS
2.3.01	March 2014	Combined previous 'Pro' and 'Gold' Editions into one.
2.2.01	September 2012	Discontinued lower record limit version. Fixed issue with trial which prevented mate recommendations from being show in some circumstances. Stats tab can now be restricted to tagged-only records; stats period is now selected by a selector within the stats tab and not through the menu.
2.1.01	July 2011	Initial public release